" A new gene called oligophrenin 1, its expression product, and the diagnostic and therapeutic applications thereof".

The present invention relates to the identification of a new gene, called oligophrenin 1, and its expression product, as well as to the diagnostic and therapeutic applications of these nucleotide and peptide sequences.

A major challenge for human genetics is the identification of new causes of mental retardation, which, although present in about 3 % of individuals, is unexplained in over half of all cases. X-linked mental retardation is acknowledged to be a major cause of severe learning difficulties. Surveys have shown an excess of males over females with severe mental retardation and later studies suggested that the excess was the result of an X-linked condition. X-linked mental retardation (XLMR) is a vastly heterogeneous group of disorders which can be roughly categorized as syndromic (MRXS) or non specific (MRX). Families with syndromic disorders usually have a quite distinct phenotypic presentation whereas families with non specific disorders present no distinctive somatic features. Despite recent advances in identifying genes such as FMR1(Verkerk et al., 1991), FRAXE (Knight et al., 1993 ; Gecz et al., 1996), L1-CAM (Vits et al., 1994), FGD1 and XH2 (Gibbons et al., 1995), involved in MRXS conditions, so far no gene significantly involved in MRX has yet been identified or cloned. Compilation of the literature and McKusick's catalogue data revealed at least 95 X-linked disorders in which mental retardation appears as the main feature. Of these 95, 40 have been regionally mapped on the X chromosome by conducting linkage studies using DNA markers in single large families or in a collection of families with the same XLMR syndrome. Several loci appear to be located in the proximal Xq region. However, it is impossible to evaluate how many MR genes there are in reality, partly because of the broad localisation and the presence of several overlaps between intervals of assignment. Thus, fine mapping and identification of genes implicated in nonspecific X-linked MR essentially depend on thorough investigation of molecular abnormalities such as balanced translocation, inversion or contiguous gene deletion associated with MR.

Bienvenu *et al.* (1997) have recently reported a molecular cytogenetic investigation of an X;12 balanced translocation observed in a female affected with a mild mental retardation and have localised the breakpoint in Xq12.

The authors of the present invention have now cloned the gene responsible for MRX, which they have called the oligophrenin 1 gene.

In order to define the genomic structure of this gene, the authors of the present invention constructed and investigated a cosmid/phage contig that covers the gene. Determination of exon-intron boundaries was performed through sequence comparison between cDNA clones and genomic DNA, which led to the identification of 25 exons.

The authors of the present invention have thus isolated and characterized the oligophrenin 1 transcripts. Said transcripts contain an open-reading frame (ORF) which is encoded by exon 2 to exon 24. This ORF is 2406 bases long and encodes a protein of 802 amino acids, called the oligophrenin 1 protein.

A subject of the present invention is thus an isolated nucleic acid having a sequence selected from the group consisting of sequence SEQ ID n° 1 to SEQ ID n° 25, and a homologous nucleotide sequence thereof.

SEQ ID n° 1 represents the 5' fragment of the genomic DNA of the human oligophrenin 1 gene.

SEQ ID n° 2 to SEQ ID n°25 represent fragments of the genomic DNA of the human oligophrenin 1 gene including exons as shown in table 1.

Table 1 : identification of the sequences

Sequence Identification	exon included in the fragment of the genomic DNA of the oligophrenin 1 gene
SEQ ID n° 2	exon 1 and exon 2
SEQ ID n° 3	exon 3
SEQ ID n°4	exon 4
SEQ ID n°5	exon 5
SEQ ID n°6	exon 6
SEQ ID n°7	exon 7
SEQ ID n°8	exon 8
SEQ ID n°9	exon 9
SEQ ID n°10	exon 10
SEQ ID n°11	exon 11
SEQ ID n°12	exon 12
SEQ ID n°13	exon 13
SEQ ID n°14	exon 14
SEQ ID n°15	exon 15
SEQ ID n°16	exon 16
SEQ ID n°17	exon 17
SEQ ID n°18	exon 18
SEQ ID n°19	exon 19
SEQ ID n°20	exon 20
SEQ ID n°21	exon 21
SEQ ID n°22	exon 22
SEQ ID n°23	exon 23
SEQ ID n°24	exon 24
SEQ ID n°25	exon 25

SEQ ID n° 26 represents the cDNA fragment corresponding to the common open-reading frame (ORF).

A subject of the present invention is also an isolated nucleic acid having a sequence selected from the group consisting of exon 1 to exon 25 as identified in the sequence listing and in table 2, and a homologous nucleotide sequence thereof.

Table 2 : identification of exon sequences

exon	from nucleotide n°	to nucleotide n°	Sequence which includes said exon
exon 1	1	634	SEQ ID n° 2
exon 2	778	935	
exon 3	403	498	SEQ ID n° 3
exon 4	483	544	SEQ ID n°4
exon 5	451	522	SEQ ID n°5
exon 6	416	517	SEQ ID n°6
exon 7	464	574	SEQ ID n°7
exon 8	244	348	SEQ ID n°8
exon 9	134	263	SEQ ID n°9
exon 10	383	483	SEQ ID n°10
exon 11	107	198	SEQ ID n°11
exon 12	211	289	SEQ ID n°12
exon 13	212	245	SEQ ID n°13
exon 14	172	234	SEQ ID n°14
exon 15	207	281	SEQ ID n°15
exon 16	270	354	SEQ ID n°16
exon 17	355	413	SEQ ID n°17
exon 18	80	185	SEQ ID n°18
exon 19	79	238	SEQ ID n°19
exon 20	230	377	SEQ ID n°20
exon 21	185	508	SEQ ID n°21
exon 22	320	485	SEQ ID n°22
exon 23	211	261	SEQ ID n°23
exon 24	115	156	SEQ ID n°24

"A homologous nucleotide sequence" is understood as meaning a sequence which differs from the sequences to which it refers by mutation, insertion, deletion or substitution of one or more bases.

Preferably, such homologous sequences show at least 70 % of homology, preferably 80 % of homology, more preferably 90 % of homology with any of sequences SEQ ID n° 1 to SEQ ID n° 26.

A polynucleotide of the invention, having a homologous sequence, hybridizes to the sequences to which it refers (any of sequences SEQ ID n° 1 to SEQ ID n° 26), preferably under stringent conditions. Parameters that define the conditions of stringency depend upon the temperature at which 50 % of annealed strands separate (T_m).

For sequences comprising more than 30 nucleotides, T_m is calculated as follows :

T_m = 81.5 + 0.41 (% G + C) + 16.6 Log (positive ion concentration) – 0.63 (% formamide) – (600/polynucleotide size in base pairs) (Sambrook et al, 1989).

For sequences comprising less than 30 nucleotides, T_m is calculated as follows :

$T_m = 4(G + C) + 2 (A + T)$.

Under appropriate stringent conditions avoiding the hybridization of non specific sequences, hybridization temperature is around from 5°C to 30°C, preferably from 5°C to 10° C below the calculated T_m, and hybridization buffer solutions that are used are preferably solutions with high ionic strength, such as an aqueous 6 X SSC solution for example.

A nucleotide sequence homologous to the open-reading from SEQ ID n° 26 means a nucleotide sequence which differs from the sequences to which it refers by mutation, insertion, deletion or substitution of one or more bases, or by the degeneracy of the genetic code so long as it codes for a polypeptide having the biological activity of oligophrenin 1 protein, as defined below.

Said homologous sequences include mammalian genes coding for the oligophrenin 1 protein, preferably of primate, cattle, sheep, swine, or rodent, as well as allelic variants.

The nucleic acid sequences of the invention are useful for the detection of an abnormality, such as a mutation, in the oligophrenin 1 gene or in the transcripts of the oligophrenin 1 gene. Such an analysis allows *in vitro* diagnosis of a neurological disorder associated with said abnormality.

A subject of the present invention is a method of *in vitro* diagnosis of a neurological disorder associated with an abnormality in the oligophrenin 1 gene or in the transcripts of the oligophrenin 1 gene, wherein one or more mutation(s), preferably inducing a modification of the expression of the oligophrenin 1 gene is detected in the oligophrenin 1 gene or in the transcripts of the oligophrenin 1 gene.

A subject of the present invention is also a nucleic acid comprising a sequence identical to SEQ ID n° 26 or to homologous sequences thereof, except for a one base deletion of the nucleotide 1578 as shown in SEQ ID n° 26.

The present invention relates to methods of *in vitro* diagnosis wherein the nucleic acid sequences of the invention or probes or primers derived thereof are used to detect aberrant synthesis or genetic abnormalities at the oligophrenin 1 gene level.

The present invention is more particularly directed to a method of *in vitro* diagnosis comprising the steps of :

- contacting a biological sample containing DNA with specific oligonucleotides permitting the amplification of all or part of the oligophrenin 1 gene, the DNA contained in the sample having being rendered accessible, where appropriate, to hybridization, and under conditions permitting a hybridization of the primers with the DNA contained in the biological sample ;
- amplifying said DNA ;
- detecting the amplification products ;
- comparing the amplified products as obtained to the amplified products obtained with a normal control biological sample, and thereby detecting a possible abnormality in the oligophrenin 1 gene.

The method of the invention can also be applied to the detection of an abnormality in the transcript of the oligophrenin 1 gene, by amplifying the mRNAs contained in a biological sample, for example by RT-PCR.

Thus another subject of the present invention is a method of in vitro diagnosis, as previously defined comprising the steps of :

- producing cDNA from mRNA contained in a biological sample ;
- contacting said cDNA with specific oligonucleotides permitting the amplification of all or part of the transcript of the oligophrenin 1 gene, under conditions permitting a hybridization of the primers with said cDNA ;
- amplifying said cDNA ;
- detecting the amplification products ;
- comparing the amplified products as obtained to the amplified products obtained with a normal control biological sample, and thereby detecting a possible abnormality in the transcript of the oligophrenin 1 gene.

This comparison of the amplified products obtained from the biological sample with the amplified products obtained with a normal biological sample can be carried out for example by specific probe hybridization, by sequencing or by restriction site analysis.

A subject of the present invention is also a nucleic acid sequence which specifically hybridizes with a nucleic acid sequence of the invention as previously defined or with their complementary sequences.

"A sequence which specifically hybridizes [...]" is understood as meaning a sequence which hybridizes with the sequences to which it refers under the conditions of high stringency (Sambrook et al, 1989). Such sequences are preferably oligonucleotides having at least 15, and more preferably at least 20 bases.

Such sequences, which are useful as primers or probes for the diagnosis methods according to the present invention may be preferably

selected from the group consisting of nucleic acid fragments of SEQ ID N° 2 to SEQ ID N° 26 as shown in table 3, or the complementary sequences thereof.

Table 3 : identification of the oligonucleotide sequences

Oligonucleotides		Sequence which includes said oligonucleotide
from nucleotide n°	to nucleotide n°	
727 958	746 977	SEQ ID n° 2
375 504	394 523	SEQ ID n° 3
418 551	437 570	SEQ ID n°4
423 553	445 574	SEQ ID n°5
388 540	407 559	SEQ ID n°6
436 584	458 603	SEQ ID n°7
219 363	239 381	SEQ ID n°8
108 336	128 355	SEQ ID n°9
361 492	380 511	SEQ ID n°10
81 223	100 242	SEQ ID n°11
188 300	207 319	SEQ ID n°12
166 259	189 278	SEQ ID n°13
133 250	152 269	SEQ ID n°14
151 293	170 315	SEQ ID n°15
221 363	244 382	SEQ ID n°16
305 438	324 457	SEQ ID n°17
25 218	44 237	SEQ ID n°18
51 252	70 271	SEQ ID n°19
206 383	225 402	SEQ ID n°20
151 511	170 530	SEQ ID n°21
294 496	313 515	SEQ ID n°22
179 271	198 291	SEQ ID n°23
65 165	84 184	SEQ ID n°24
152 586 641	172 606 663	SEQ ID n° 26

One skilled in the art very well knows the standard methods for analysing the DNA contained in a biological sample and for diagnosing a genetic disorder. Many strategies for genotypic analysis are available (Antonarakis et al., 1989, Cooper et al., 1991).

Preferably, one can use the DGGE method (Denaturing Gradient Gel Electrophoresis), or the SSCP method (Single Strand Conformation Polymorphism) for detecting an abnormality in the oligophrenin 1 gene. Such methods are preferably followed by direct sequencing. The RT-PCR method may be advantageously used for detecting abnormalities in the oligophrenin 1 transcript, as it allows to visualize the consequences of a splicing mutation such as exon skipping or aberrant splicing due to the activation of a cryptic site. This method is preferably followed by direct sequencing as well. The more recently developed technique using DNA chip can also be advantageously implemented for detecting an abnormality in the oligophrenin 1 gene (Bellis et al., 1997).

The cloning of the oligophrenin 1 gene, as well as the identification of various mutations responsible for neurological disorders according to the invention, allow direct or semi-direct diagnosis. The specificity and reliability of such diagnosis methods are more particularly appreciable for prenatal diagnosis. The nucleic acid sequences of the present invention represent a highly interesting tool for genetic counseling.

The authors of the present invention have shown that the oligophrenin 1 protein is a rho-GAP protein and that the constitutional loss of oligophrenin 1 activity in humans results in cognitive impairment. Defects in the oligophrenin 1 gene, or in the oligophrenin 1 gene product may cause inactivation of the oligophrenin 1 protein, which leads to constitutive activation of its target GTPases.

Such constitutive activation of rho family members has been shown to produce marked changes in certain actin-based processes, to alter the cytoskeleton in cultured cells (Nobes et al., 1995) and to affect cell migration and axon outgrowth in vivo (Luo et al., 1994). In addition,

constitutively active Rac1 rho-GTPase produced neuropathological changes in transgenic mice such as defects in axon outgrowth and dendritic spine morphogenesis in mouse Purkinje cells (Luo et al., 1996).

In the same manner, constitutive activation of a rho-GTPase resulting from the loss-of-function of oligophrenin 1 protein leads to a dysfunctioning of signal transduction pathways involved in cell migration and axon outgrowth during development of the nervous system. Mental retardation may be the clinical expression of such neuropathological changes.

The oligophrenin 1 gene would thus be involved in disorders due to an abnormal neurone migration. Such disorders include not only genetic disorders such as nonspecific X-linked mental retardation but also incurable cryptogenic epilepsies and neurodegenerative diseases, such as Alzheimer's disease and cognitive impairments related to aging.

The present invention also provides transgenic non-human animals or cells thereof.

Said transgenic animal can have a exogenous oligophrenin 1 protein of this invention due to the presence of a gene encoding and expressing that protein or part of that protein.

Transgenic animals are generally well known, as are their methods of production.

The present invention contemplates a non-human animal containing a oligophrenin 1 gene of the present invention integrated in the genome of the animal's somatic and germ cells, i.e., a transgenic animal, preferably transgenic mammals.

Animals containing a transgene encoding a oligophrenin 1 protein of the present invention are typically prepared using the standard transgenic technology described in Hogan et al. (1987) and Palmiter et al. (1986). Production of transgenic mammals is also possible using the homologous recombination transgenic systems described by Capecchi (1989). Preparation of transgenic mammals has also been described in WO94/21670.

One technique for transgenically altering a mammal is to microinject a rDNA into the male pronucleus of the fertilized mammalian egg to cause one or more copies of the rDNA to be retained in the cells of the developing mammal.

Alternative methods for producing a non-human mammal containing a rDNA of the present invention include infection of fertilized eggs, embryo-derived stem cells, totipotent embryonal carcinoma (Ec) cells, or early cleavage embryos with viral expression vectors containing the rDNA (Palmiter et al., (1986)).

A transgenic animal can also have a mutation in its own native oligophrenin 1 gene, thereby rendering the oligophrenin 1 protein non-functional (i.e., a "knockout" trangenic animal). Such an animal is useful as it presents the clinical conditions associated with the defects in the mutated oligophrenin 1 protein, and further can be a model for evaluation of candidate therapeutics that would treat subjects with defects in that protein.

More particularly, transgenic non-human animals or cells in culture, that overexpress oligophrenin 1 protein or preferably express a native oligophrenin 1 protein that has been rendered non-functional ("knock-out" transgenic animal) may be useful in a method for screening chemical entities or drugs likely to act on the signaling pathway to which the rho-GAP^{MRX} protein (oligophrenin) belongs.

Transgenic non-human animals or cells thereof that overexpress oligophrenin 1 protein refer to animals or cells thereof that express an exogenous oligophrenin 1 protein of the invention in addition not the native protein.

In one embodiment, the screening method of the invention comprises the steps of :

i) administering a drug to be tested to a transgenic non-human animal that overexpress oligophrenin 1 protein or preferably express a native oligophrenin 1 protein that has been rendered non-functional ; and

ii) observing clinical expression of neuronal changes *in vivo* and/or *in vitro* culturing nervous cells from said animal and observing the stimulation or recovery of axon outgrowth or morphogenesis.

In another embodiment, the screening method of the invention comprises the steps of :

i) contacting a drug to be tested with nervous cells or nervous tissue cultures that overexpress oligophrenin 1 protein or preferably express a native oligophrenin 1 protein that has been rendered non-functional ; and

ii) measuring the axon outgrowth.

In the above embodiment, said cells are either obtained from said transgenic animals or are established cell lines, such as neuroblastoma or primary cultures of neuronal cells, which have been transfected by a DNA construct, e.g by means of a viral vector, allowing the expression of exogenous oligophrenin 1 protein or rendering the native oligophrenin 1 non functional.

Drugs selected by the methods of screening as above-defined, and pharmaceutical compositions containing such a drug in association with a pharmaceutically acceptable carrier, are also encompassed by the present invention.

The ORF of the oligophrenin 1 gene as shown in SEQ ID n° 26 according to the invention encodes a protein of 802 amino acids with a relative molecular mass of 91 kD. Hydropathy analysis (Kyte and Doolittle, 1982) suggests that the oligophrenin 1 protein is hydrophilic. Based on consensus motifs in PROSITE database (Bairoch et al., 1997), many potential phosphorylation sites were predicted including a tyrosine kinase phosphorylation site at position 142. Comparison of the protein sequence with other sequences in the databases indicated that the oligophrenin 1 gene encodes a rho-GAP containing protein.

Sequence alignment shown in figure 3b illustrates the remarkable similarity between the predicted oligophrenin 1 domain and other members of the rho-GAP subfamily. This similarity extends over 180 residue region localised in the central part of the predicted protein and concerns the three

structurally conserved regions (SCRs) that are specific to the rho-GAP proteins.

Among rho-GAPs, the oligophrenin 1 protein showed the greatest similarity to the chicken Graf protein (Hildebrand et al., 1996). This similarity extends on both sides of the rho-GAP domains, but oligophrenin 1 does not contain the SH3 domain reported for the Graf protein. The rho-GAP activity of the oligophrenin 1 protein is consistent with the functional analysis of the chicken Graf protein, which has both part of the N-terminal and rho-GAP domains identified in the oligophrenin 1 protein. Graf protein has been shown to preferentially stimulate the GTPase activity of the GTP-binding proteins RhoA and Cdc42 (Hildebrand et al., 1996).

The C-terminal part of oligophrenin 1 protein does not match any known sequence, whereas the N-terminal domain of oligophrenin 1 protein is similar to a highly conserved protein, of unknown function, identified in *C. elegans*, mouse and human (Fig. 3c). This protein presents two isoforms identified as CELZK328 and CELT04C95 (Genbank), which correspond to two different ORF.

A further subject of the present invention is thus an isolated oligophrenin 1 polypeptide substantially comprising the aminoacid sequence of SEQ ID n° 27 or a homologous aminoacid sequence thereof.

The above term "substantially" is understood as meaning that said isolated oligophrenin 1 polypeptide exhibits the same biological and/or immunological properties, as the native oligophrenin 1 protein.

More particularly said aminoacid sequence may be SEQ ID n° 26, or a homologous aminoacid sequence thereof.

"A homologous aminoacid sequence" is understood as meaning a sequence which differs from the sequences to which it refers by mutation, insertion, deletion or substitution of one or more aminoacids, without inducing modification of biological and/or immunological properties. Said derivative

aminoacid sequence shows at least 60 % of homology, preferably 70 % of homology, preferably 80 % of homology with the oligophrenin 1 polypeptide of SEQ ID n° 26.

The "biological properties" of the polypeptides of the invention refer to the activity of the oligophrenin 1 protein, which enhances GTPase activity of small Ras-like GTPases and hence turns them off.

The "immunological properties" of the polypeptides of the invention refer to the ability of the polypeptides of the invention to induce an immunological response mediated by antibodies which recognize the oligophrenin 1 polypeptide of the invention.

The polypeptides according to the invention can be obtained by any of the standard methods of purification of soluble proteins, by peptide synthesis or by genetic engineering. Said techniques comprise the insertion of a nucleic acid sequence coding for a peptide of the invention into an expression vector, such as a plasmid, and the transformation of host cells with the expression vector, by any of the methods available to the skilled person, like for instance electroporation.

The present invention thus relates to vectors for cloning and/or expression comprising a nucleic acid sequence of the invention and to host cell transfected with these vectors. The expression vector according to the invention comprises a nucleic acid sequence encoding a polypeptide of the invention, said nucleic acid sequence being operably linked to elements allowing its expression. Said vector advantageously contains a promoter sequence, signals for initiation and termination of translation, as well as appropriate regions for regulation of translation. Its insertion into the host cell may be transient or stable. Said vector may also contain specific signals for secretion of the translated protein.

These various control signals are selected according to the host cell which may be inserted into vectors which self-replicate in the selected host cell, or into vectors which integrate the genome of said host.

Host cells may be prokaryotic or eukaryotic, including but not limiting to bacteria, yeasts, insect cells, mammalian cells, including cell lines which are commercially available.

A subject of the present invention is also a method for producing a recombining oligophrenin 1 polypeptide, wherein said host cell is transfected with said expression vector and is cultured in conditions allowing the expression of a polypeptide according to the invention.

The present invention also relates to monoclonal or polyclonal antibodies, or fragments thereof, or chimeric or immunoconjugate antibobies, which are capable of specifically recognizing a polypeptide according to the invention.

Polyclonal antibodies can be obtained from serum of an animal immunized against the oligophrenin 1, which can be produced by genetic engineering for example, as above described, according to standard methods well-known by one skilled in the art.

Monoclonal antibodies can be obtained according to the standard method of hybridoma culture (Kohler and Milstein, 1975).

The antibodies of the present invention can be chimeric antibodies, humanized antibodies, or antigen binding fragments Fab and F(ab')2. They can also be immunoconjugated or labelled antibodies.

Said antibodies are particularly useful for detecting or purifying a oligophrenin 1 polypeptide according to the invention in a biological sample.

They are more particularly useful for detecting an abnormal expression of the oligophrenin 1 protein in connection with neurological disorders as above described.

Another subject of the present invention is a pharmaceutical composition comprising a purified oligophrenin 1 polypeptide of the invention and/or a homologous polypeptide thereof, in association with a pharmaceutically acceptable carrier.

A further subject of the present invention is a pharmaceutical composition comprising a nucleic acid encoding said polypeptides and a pharmaceutically acceptable carrier. Said nucleic acid, preferably inserted in a vector, may be administered in a naked form or in association with transfection facilitating agents.

A further subject of the invention is a pharmaceutical composition comprising an anti-sense sequence capable of specifically hybridizing with a nucleic acid sequence encoding said polypeptides, in association with a pharmaceutically acceptable carrier.

A still further subject of the invention is a pharmaceutical composition comprising an antibody directed against said polypeptides, in association with a pharmaceutically acceptable carrier.

Preferably the present invention is directed to a pharmaceutical composition comprising a purified oligophrenin 1 polypeptide of the invention and/or a homologous polypeptide thereof, in association with a pharmaceutically acceptable carrier.

The term "homologous polypeptide", as active ingredient of a pharmaceutical composition, refers to a polypeptide with a homology of at least 40 %, preferably of at least 60 % in comparison to the oligophrenin 1 protein. Such homologous polypeptides include any known protein which exhibits a rho-GAP activity. Preferably said homologous polypeptide is for example the protein CELZK328 or CELT04C95.

The pharmaceutical compositions of the invention are useful for preventing and/or treating neurological disorders, wherein the oligophrenin 1 protein or a homologous protein thereof is implicated. As above underlined, the authors of the present invention have shown that defects in a Ras-like GTPase (rho-GAP) dependent signalling pathway are associated with cognitive impairment, resulting from misguided axon growth and/or defective cell migration. Consequently, the disorders which are more particularly aimed at are disorders of the central nervous system in connection with the axonal

development, more particularly a disorder associated with cognitive impairment. Such disorders include nonspecific X-linked mental retardation, as well as cryptogenic epilepsies or neurodegenerative diseases, such as Alzheimer's disease and cognitive impairments related to aging.

Another subject of the invention is the use of a purified oligophrenin 1 polypeptide of the invention and/or a homologous polypeptide thereof, in association with a pharmaceutically acceptable carrier for the manufacture of a medecine for preventing and/or treating neurological disorders, wherein the oligophrenin 1 protein or a homologous protein thereof is implicated.

The pharmaceutical compositions of the invention may be administered to a mammal, preferably to a human, in need of a such treatment, according to a dosage which may vary widely as a function of the age, weight and state of health of the patient, the nature and severity of the complaint and the route of administration.

The appropriate unit forms of administration comprise oral forms such as tablets, gelatin capsules, powders, granules and oral suspensions or solutions, sublingual and buccal administration forms, subcutaneous, intramuscular, intravenous, intranasal or intraoccular administration forms and rectal administration forms.

A further subject of the present invention is a method of preventing and/or treating neurological disorders resulting from defects in the oligophrenin 1 gene or in the oligophrenin 1 protein or in a homologous gene or protein thereof, which comprises administering to a subject in need of a such treatment an amount of a pharmaceutical composition as above defined effective to prevent and/or alleviate said neurological disorders.

The present invention is further illustrated by, but not limited to, the figures and the examples that follow.

LEGENDS TO FIGURES :

Figure 1. (**a**) Physical map of the Xq12 locus and genomic structure of the oligophrenin 1 gene spanning the X-chromosomal breakpoint. YAC, PAC and cosmid contigs are indicated with lines and STSs by vertical bars. The STS C16T3 was generated from the distal end of the 9 kb HindIII fragment containing the X-chromosomal breakpoint (represented by an arrow). The oligophrenin 1 gene spans at least 300 kb and it consists of 25 exons of which 23 are coding (white boxes). (**b**) Southern blot analysis of HindIII digested genomic DNAs from the patient exhibiting the (X;12) translocation and a normal female using the STS C16 T3 as probe. The junction fragment is indicated by JF.

Figure 2. Fetal and multiple adult tissue northern blots containing poly(A)$^+$ RNA (Clontech®) hybridized with C2 cDNA probe. A 7.5 kb transcript was observed after an overnight exposure at - 80°C. Hybridization of Northern blots with an actin probe was performed to assess differences between amounts of loaded poly(A)$^+$ RNA samples.

Figure 3. (**a**) Coding part of the cDNA and deduced amino acid sequences of the oligophrenin 1 gene. Nucleotides in bold correspond to exon-exon boundaries. The GAP domain is underlined, Primers used in RT-PCR to study the expression of the gene in the patient with the X;12 translocation are double underlined and the internal primer used for hybridization to ascertain PCR products is shown in dotted line. The deleted nucleotide at position 1578 is indicated in italic bold case. (**b**) Sequence alignment of the GAP domain present in oligophrenin 1 and different rhoGAP proteins reviewed in Aelst et al., 1997. The GAP domain contains three Structurally Conserved Regions (SCRs) (Boguski et al., 1993). CELZK328 corresponds to an ORF predicted from sequences of the C.*elegans* genome. (**c**). Sequence alignment showing

the high conservation of oligophrenin 1 N-terminal domain. CELTO4C95 corresponds to another C. elegans ORF, different from CELZK328. EST483210 is part of the mouse homologue oligophrenin 1 cDNA and EST387042 corresponds to a human EST that was localized on chromosome 11 by the Whitehead Institute.

Identical residues are indicated by black shading, similar residues by grey. The alignment was performed using Multalin (Corpet et al., 1988) and Boxshade softwares.

Figure 4. (**a**). Study of oligophrenin 1 transcript in the patient with the (X;12) translocation and in a normal female used as control (XX). Southern blot of PCR products amplified from total RNAs isolated from EBV transformed lymphoblastoid cell lines.

(**b**) Nucleotide sequences showing the one base pair deletion. Direct sequencing of PCR products corresponding to exon 19 of IV-3 proband.

(**c**). Segregation analysis of the mutated allele in XLMR family D. Open squares, unaffected males; closed squares, affected males; open circles, unaffected females; dotted symbols, phenotypically normal carrier females.

EXAMPLE 1 : Identification of the oligophrenin 1 gene :

1. Experimental procedures

Case report and family materials

Clinical data and diagnosis concerning the female patient with the t(X;12) translocation were previously described (Luo et al., 1996). Concerning the MRX families, a linkage study was reported in The European XLMR Consortium 's report (1997).

YAC and PAC clones

YAC clones of the Xq12 locus were obtained from the UK HGMP Resource Centre. PAC clones were obtained from the German resource center (RZPD). Primer sequences corresponding to STSs are available in Genome Data Base. STSC16T3 is a 189 bp fragment amplified with the primers: C16T3F (5' CACAGCAAGCAATAAGCACT 3') and C16T3R (5' TGTCTCCTGTGCTCTTTCCA 3'). Overlaps between clones and STS mapping were performed by a combination of STS/EST amplification and hybridization approaches.

cDNA isolation

Approximately 1 X 10^6 recombinant clones of a lgt10 human fetal brain cDNA library (Clontech) were plated and screened following standard techniques (Sambrook et al., 1989). Library screening was performed using as probe RT-PCR products generated with primers located within the predicted exons. Positive clones were plaque purified and their inserts were subcloned into bluescript vector and sequenced. 3' RACE PCR (Clontech kit) was used to obtain the 3' end of the cDNA. The full-length cDNA is a composite of 14 clones (only 4 clones are represented on figure 1a).

Genomic DNA sequencing of human cosmid clones:

Human cosmid clones were detected in the Imperial Cancer Research Fund (ICRF) flow-sorted human X chromosome library (Nizetic et al., 1991) using B2 cDNA as a probe. Cosmid 2C6, 4D2 and 35 shown in figure 1a are from a cosmid library corresponding to the YAC 4690. Exon-intron boundaries were identified through sequence comparisons between cDNA and genomic DNA clones. To generate genomic sequences, DNAs of cosmid clones were used as templates and primed with exonic oligonucleotides. ICRF coordinates of cosmid clones shown in figure 1 are as follow: cos12: ICRFc104J1515Q8, cos15: ICRFc104K1628Q8, cos7: ICRFc104P0212Q8, cos3: ICRFc104B1719Q8, cos11: ICRFc104F178Q8, cos5: ICRFc104B1515Q8.

2. Results :

As a first step in identifying a potential gene in Xq12 involved in MRX, Bienvenu *et al.* (1997) have reported a molecular cytogenetic investigation of an X;12 balanced translocation observed in a female affected with a mild mental retardation. The Xq breakpoint was localized within an ICRF 850 kb YAC clone y900H0493 (4690), positive for PGKP1 and DXS159 markers.

The authors of the present invention have used long-range restriction maps of the YAC clone 4690, FISH analyses and somatic hybrid cell lines containing the derivative chromosome 12 as their only human X-chromosomal component to regionally fine map the X chromosomal breakpoint. Figure 1a depicts the location of the translocation breakpoint on the normal X and summarizes YAC, PAC and cosmid contigs spanning the breakpoint. The probe, STSC16T3, which detects the junction fragment and localized the breakpoint to a 9 kb HindIII fragment (fig 1b) was isolated from the cosmid clone 4D2 (Fig 1a). Aberrant bands confirming the latter results

were also obtained by hybridization of the same probe to a Southern blot containing DNA from the patient digested with several other enzymes.

Sequencing of randomly subcloned HindIII fragments including the 9 kb fragment isolated from cosmid clones spanning the breakpoint, and searches for homology in data bases revealed sequence identities between the isolated sequences and those corresponding to the PAC clone 360E18 (Fig 1a) which were generated by the Sanger Centre (Cambridge,UK). Available sequences were then used for computational analyses and comparison with nucleotide and protein sequences. Some of the potential exons identified by GRAIL (Kel et al., 1993) and FEXH/HEXON (Lerman et al., 1987) programs showed a significant homology with the mouse EST 483210. The predicted polypeptide corresponding to this EST revealed a significant homology with the human EST 387042 localised on chromosome 11 and with the *C. elegans* ORF CELT04C95. Further investigations suggested that the *C. elegans* ORF is represented on the genomic sequences derived from the PAC clone by 8 different potential exons scattered over 130 kb (Fig. 1a). These predictions were confirmed by RT-PCR experiments using primers located within the potential exons and human fetal brain total RNA. Furthermore, hybridization of the RT-PCR products to a Northern blot containing polyA$^+$ RNA detected a 7.5 kb transcript most highly expressed in fetal brain (Fig. 2). Together, these results indicated the presence of a candidate gene located in the vicinity of the translocation breakpoint.

In order to identify the full length cDNA the authors of the present invention used a combination of fetal brain cDNA library screening, PCR and rapid amplifications of cDNA ends (RACE). This approach enabled to obtain a composite full length nucleotide sequence of the cDNA (Figure 3a).

To confirm that the identified gene is disrupted by the translocation breakpoint the structure of the gene including the exon-intron boundaries was determined through sequence comparisons between cDNA and cosmid genomic DNA clones isolated either from a cosmid library generated from YAC clone 4690 or the ICRF flow sorted X-specific cosmid

library (fig 1a). The physical mapping of the 25 exons allowed to demonstrate that the candidate MRX gene is transcribed from telomere to centromere and the translocation breakpoint maps within the second intron leading therefore to a displacement on the derivative chromosome 12 of the first two exons including the one containing the putative translation initiation codon. Confirmation of this latter result was obtained by FISH analysis using as probe a cosmid clone containing the first two exons of the gene (cos 12, fig 1a). This experiment showed that this cosmid maps exclusively on the derivative chromosome 12 (data not shown).

To investigate the gene expression and examine whether both alleles are inactive, RT-PCR experiments were performed on RNA isolated from EBV-transformed lymphoblastoid cell lines (LCL) of the patient and a female control using primers located in exons 1 and 2. These experiments failed to amplify a normal gene product on RNA of the patient (fig 4a). Consistently, the normal X chromosome was found to be late replicating, indicating preferential X-inactivation of the normal X chromosome (Bienvenu et al., 1997), and the MRX candidate gene was found to undergo X-inactivation.

EXAMPLE 2 : Identification of mutations in MRX families

1. Experimental procedure

Mutations analysis

Genomic DNA was extracted from EBV - transformed lymphoblastoid cell lines using standard protocols. The 23 coding exons and exon-intron boundaries were individually amplified with specific primers. In each amplification one primer was a 5' psoralen-modified primer (Fernandez et al., 1993). PCR products were checked on standard agarose gels prior to analysis by the DGGE technique. When an aberrant pattern of migration was observed, the corresponding PCR product was purified and directly sequenced

on both strands using the Dye Terminator Cycle Sequencing kit protocol (Applied Biosystems).

Figure 4(a) represents a Southern blot of PCR products amplified from total RNAs isolated from EBV transformed lymphoblastoid cell lines. RT-PCR were performed with (+) or without (-) reverse transcriptase and cDNAs were amplified for 40 cycles with primers located in exon 1 and 2 (figure 3, double underlined nucleotide sequences). After gel electrophoresis, the Southern blot was hybridized with an internal oligonucleotide (dotted line on figure 3). The 650 bp fragment corresponds to an amplification from the contaminating genomic DNA (intron 1 is 140 bp long). The lane labelled DNA, corresponds to the PCR product obtained from a female total genomic DNA. The negative control indicated by Ct(-) corresponds to a PCR experiment without template. RT-PCR amplification of the ubiquitously expressed transcript produced by the distal part of the dystrophin gene was used as internal standard.

The mutation to co-segregate with the mental retardation phenotype as shown on figure 4c was detected by denaturing gradient gel electrophoresis of PCR products corresponding to exon 19 of the oligophrenin 1 gene. Exon 19 was amplified by PCR with primers 19F (5' GTT AAT CTT GCC CCT TTT CT 3') and 19R (5' Psoralen- TA GGA AGA CAG GTA GTG AGA AT) yielding a 221 bp product. 10 µl of each amplified product was mixed with 10 µl of normal control PCR product. Heteroduplexes were generated by denaturing for 10 min, and subsequent reannealing for 45 min at 56°C. The samples were electrophoresed through a 25-65% denaturant 6% polyacrylamide gel for 7.5h at 60°C and 160V. The characteristic shifted profile displayed by the mutated allele allow an easy study of the familial segregation.

2. Results

In order to prove that the isolated gene is responsible for non-specific mental retardation, four unrelated probands from MRX families (The

European XLMR Consortium, 1997), previously mapped in genetic intervals which encompass the candidate gene, were analysed for the presence of point mutations. The strategy involved investigation by DGGE (Lerman et al., 1993) (denaturing gradient gel electrophoresis) of PCR products corresponding to all coding exons and sequencing of exons exhibiting abnormal migration profiles. PCR primers were designed not only to amplify individual exons but also sequences flanking the exons. DGGE analyses of amplified products, corresponding to exon 19, from the proband IV-3 of the family D (fig 4c) showed an abnormal shift in mobility. Compared with the normal product, sequence analysis showed that the aberrant product contained a one base pair deletion of the nucleotide 1578 (fig 4b); the resulting frameshift mutation was predicted to cause premature translation termination four codons downstream of the mutation. Cosegregation of the mutation with the disease which was confirmed in the large family using the DGGE technique (fig 2b) and the absence of this mutation in 100 control individuals indicate that the deletion does indeed cause the mutant phenotype.

EXAMPLE 3 : Expression of the oligophrenin 1 gene

C2 cDNA clone was isolated from a fetal brain cDNA library. C2 DNA probe was used to hybridize poly(A)$^+$ RNA (Clontech©) contained in fetal and multiple adult tissues. A 7.5 kb transcript was observed after an overnight exposure at - 80°C. Hybridization of Northern blots with an actin probe was performed to assess differences between amounts of loaded poly(A)$^+$ RNA samples.

As shown in figure 2, the oligophrenin 1 transcript was most abundant in RNA from fetal brain. A lower level of expression was also detected in several other tissues including adult brain. To further investigate the distribution of the transcript during development, *in situ* hybridization was used to examine the expression of the mouse homologous gene in embryonic days (E) 10.5, E12.5, E14, E18 and in postnatal day 1 of mouse embryos and

postnatal tissues. In addition to a low expression in all tissues with no significant differences, it was found that the gene is expressed at a higher level in all parts of the developing neuroepithelium of the neural tube. During later differentiation stages and in the mature brain a significant level of expression is visible in different structures of the brain with no striking distribution of the mRNA expression.

Several lines of evidence show that defects in oligophrenin 1 are responsible for X-linked non-specific mental retardation. First, the oligophrenin 1 gene maps to a potential mental retardation genetic locus in Xq12 identified by linkage analyses (Lubs et al., 1996, The European XLMR Consortium, 1997). Second, literature reports (Davies, 1997) concerning two patients with complete androgen insensitivity syndrome (CAIS) and mental retardation showed the presence of deletions which include several markers both proximal and distal to the AR gene and extend to DXS905 and DXS908. The above mapping data showed that these markers are located within the second and fifth intron of the oligophrenin 1 gene (Fig. 1a), demonstrating therefore that most of the exons are deleted in these two patients with mental retardation. In contrast, deletions in two CAIS patients without mental retardation do not extend to the oligophrenin 1 gene as deletions are limited to the androgen receptor gene itself (Davies, 1997). Third, investigation of this gene in the female patient with mental retardation and a t(X;12) demonstrated an absence of expression of both alleles resulting from the disruption of one allele by the translocation breakpoint and a preferential inactivation of the second allele. Fourth, a one base pair deletion within the rho-GAP domain, predicted to result in a severe abbreviated and nonfunctional oligophrenin 1 protein, cosegregates with a recessive mental retardation phenotype in a large affected family mapped to the pericentromeric region. Finally, the oligophrenin 1 mRNA is highly expressed in fetal brain, a finding consistent with the disease phenotype.